Title	条件数制約つき正定値行列近似問題について (最適化手法の理論と応用の繋がり)
Author(s)	田中, 未来; 中田, 和秀
Citation	数理解析研究所講究録 (2013), 1829: 113-121
Issue Date	2013-03
URL	http://hdl.handle.net/2433/194804
Right	
Type	Departmental Bulletin Paper
Textversion	publisher

数理解析研究所講究録
第 1829 巻 2013 年 113-121

条件数制約つき正定値行列近似問題について*

田中未来 (Mirai Tanaka)† 中田和秀 (Kazuhide Nakata)‡

概要

本論文では条件数制約つき正定値行列近似問題について考察する. 条件数制約つき正定値行列近似問題とは, 条件数がある値以下の正定値行列の中で与えられた対称行列に最も近いものを求める最適化問題である. 本論文では, 計量としてユニタリ相似不変ノルムを用いる場合はこの問題を簡単な構造をもつ問題に変形できることを示す. 特に Ky Fan の p-k ノルムを用いた場合は 1 変数の区分凸最適化問題に帰着できる. さらに, スペクトルノルムとトレースノルムを用いる場合における解析解を示す.

1 はじめに

確率変数 $\boldsymbol{x} \in \mathbb{R}^n$ の分散共分散行列 $\boldsymbol{X}$ を観測データ $\boldsymbol{x}_1, \dots, \boldsymbol{x}_N$ から推定する状況を考える. 分散共分散行列 $\boldsymbol{X}$ の最尤推定量 $\widehat{\boldsymbol{X}}$ は次のように表される:

$$\widehat{\boldsymbol{X}} = \frac{1}{N} \sum_{i=1}^{N} (\boldsymbol{x}_i - \overline{\boldsymbol{x}})(\boldsymbol{x}_i - \overline{\boldsymbol{x}})^{\mathrm{T}}. \tag{1}$$

ここで $\overline{\boldsymbol{x}} = (1/N) \sum_{i=1}^{N} \boldsymbol{x}_i$ である. これに対して金融の実務などにおいては, アナリストの分析などによる事前情報が $X_{ij} \geq 0$ であるにもかかわらず $\widehat{X}_{ij} < 0$ となってしまった場合に推定値を事後的に $\widehat{X}_{ij} = \widehat{X}_{ji} = 0$ と補正するなどということがしばしば行なわれるということが [1] で述べられている. この場合, 結果として得られる推定値 $\widehat{\boldsymbol{X}}$ は必ずしも半正定値性を満たさないため, 分散共分散行列として不適切なものとなりうる. このような状況において, 半正定値でない行列 $\widehat{\boldsymbol{X}}$ をなんらかの意味で最も近い半正定値行列 $\boldsymbol{X}$ で近似することは自然な発想だろう. これは次のような最適化問題として定式化できる:

$$\left| \begin{array}{ll} \text{minimize} & \|\boldsymbol{X} - \widehat{\boldsymbol{X}}\| \\ \text{subject to} & \boldsymbol{X} \in \mathcal{S}^n_+. \end{array} \right. \tag{2}$$

ここで $\|\cdot\|$ はなんらかの行列ノルムであり, $\mathcal{S}^n_+$ は n 次半正定値行列のなす錐を表す. この問題は行列ノルムとして Frobenius ノルムを用いる場合, 行列 $\widehat{\boldsymbol{X}}$ の固有値分解を $\sum_{i=1}^{n} \lambda_i \boldsymbol{p}_i \boldsymbol{p}_i^{\mathrm{T}}$ とすれば, $\boldsymbol{X}^* = \sum_{i:\lambda_i>0} \lambda_i \boldsymbol{p}_i \boldsymbol{p}_i^{\mathrm{T}}$ が最適解となる.

---

* 本研究は科研費 (22710136) の助成を受けたものである.

† 東京工業大学 大学院社会理工学研究科 経営工学専攻 (Department of Industrial Engineering and Management, Graduate School of Decision Science and Technology, Tokyo Institute of Technology)

‡ 東京工業大学 大学院社会理工学研究科 経営工学専攻 (Department of Industrial Engineering and Management, Graduate School of Decision Science and Technology Tokyo Institute of Technology)

しかしながら, この方法によって得られる X^* は $\widehat{X}$ が正定値でない限り正則でないため, 逆行列が存在しない, あるいは Cholesky 分解ができないなどといった不都合が後の分析において生じる. 同様の問題が, 高次元小標本データからの推定においても起こる. 高次元小標本データからの推定の場合, 式 (1) で表される最尤推定量 $\widehat{X}$ について $\mathrm{rank}(\widehat{X}) \leq N \ll n$ であるため, $\widehat{X}$ は正則でない.

このような推定量の不安定性については統計的な観点から研究が行なわれてきた [2,3,4,7,8,9]. 特に Won and Kim [9] は, 半正定値制約と条件数制約の下で尤度を最大化する推定方法を提案している. ここで 行列 X の条件数 $\mathrm{cond}(X)$ とは, 行列 X の最大特異値 $\sigma_{\max}(X)$ と最小特異値 $\sigma_{\min}(X)$ の比 $\sigma_{\max}(X)/\sigma_{\min}(X)$ であり, 行列の数値的な扱いやすさを表す標準的な指標である. Won and Kim [9] はこの推定量を得るためには 1 変数の最適化問題を解けばよいということを示している.

しかしながら, この方法は尤度の最大化に基づくため, 不定値な行列の補正に用いることはできない. そこで, 本論文では問題 (2) に対して条件数制約を付加した次のような問題を考える:

$$\left|\begin{array}{ll} \text{minimize} & \|X - \widehat{X}\| \\ \text{subject to} & X \in \mathcal{S}^n_+, \\ & \mathrm{cond}(X) \leq \kappa. \end{array}\right. \tag{3}$$

この方法は尤度の最大化に基づくものではないため, 最尤推定以外の推定をはじめとする幅広い分野における不定値な行列の補正に応用できる可能性がある. なお, この問題の場合は行列 X の半正定値性より, 行列 X の最大固有値 $\lambda_{\max}(X)$ と最小固有値 $\lambda_{\min}(X)$ を用いて $\mathrm{cond}(X) = \lambda_{\max}(X)/\lambda_{\min}(X)$ と表すことができる. 本論文ではこの問題とほぼ等価な次の問題を条件数制約つき正定値行列近似問題と呼び, この問題について考察する:

$$\left|\begin{array}{ll} \text{minimize} & \|X - \widehat{X}\| \\ \text{subject to} & X \in \mathcal{S}^n_+, \\ & \lambda_{\max}(X) \leq \kappa\lambda_{\min}(X). \end{array}\right. \tag{4}$$

問題 (3) と問題 (4) の差は, 後者の問題において $X = O$ を実行可能とする点のみである. この問題は目的関数における行列ノルムとして適当なノルムを用いると対称錐最適化問題に帰着できるが, 計算量の観点から考えるとこの方法で大規模な問題を解くことは難しい.

本論文の第 2 節では, 問題 (4) の目的関数における行列ノルムとしてユニタリ相似不変ノルムを用いる場合, 問題 (4) を簡単な構造をもつ問題に変換できることを示す. ここで $\mathbb{C}^{n\times n}$ 上の行列ノルム $\|\cdot\|$ がユニタリ相似不変であるとは, 任意の $X \in \mathbb{C}^{n\times n}$ と任意のユニタリ行列 $U \in \mathbb{C}^{n\times n}$ に対して,

$$\|UXU^{\mathrm{H}}\| = \|X\|$$

が成り立つことを意味する [6]. ここで $U \in \mathbb{C}^{m\times n}$ に対して U^{H} は U の共役転置を表す. さらに第 2 節では, ユニタリ不変ノルムとして Ky Fan の p-k ノルムを用いる場合, 問題 (4) を 1 変数の区分凸最適化問題に帰着できることを示す. ここで $X \in \mathbb{C}^{m\times n}$ の Ky Fan の p-k ノルムとは, パラメータ $p \geq 1, k \in \{1, \ldots, \min\{m, n\}\}$ を用いて次のように定義されるノルムである [6]:

$$\|X\|_{p,k} = \left(\sum_{i=1}^{k} \sigma_i(X)^p\right)^{1/p}.$$

ここで $\sigma(X)$ は X の重複を許して i 番目に大きい特異値を表す. このノルムがユニタリ相似不変であることは特異値がユニタリ相似不変であることからすぐにわかる. このノルムは $p = 1$ のとき, 単に Ky Fan の k ノルムと呼ばれ, $k = n$ のとき Schatten p ノルムと呼ばれる. 特に, $k = 1$ のときスペクトルノルム (誘導 2 ノルム), $p = 1, k = n$ のときトレースノルム, $p = 2, k = n$ のとき Frobenius ノルムとそれぞれ呼ばれる. 第 3 節では, スペクトルノルムとトレースノルムを用いる場合について, 帰着した 1 変数の区分凸最適化問題の解析解を示す. 第 4 節ではまとめと今後の課題を述べる.

なお, 以下では $\mathcal{H}^n$ で n 次 Hermite 行列の空間を表す. また, $X \in \mathcal{H}^n$ に対して $\mathrm{diag}(X)$ は X の対角成分を並べたベクトルを, $x \in \mathbb{C}^n$ に対して $\mathrm{Diag}(x)$ は v の各成分を対角成分に並べた対角ベクトルを表す. さらに, $X, Y \in \mathbb{C}^{m \times n}$ に対して $Z = X \circ Y$ は Hadamard 積を表す. すなわち $Z_{ij} = X_{ij}Y_{ij}$ である.

2 解きやすい定式化

本節では問題 (4) の計量にユニタリ不変ノルムを用いる場合に問題の構造を簡単にすることができることを示し, 特殊な場合として Ky Fan の p-k ノルムを用いる場合には 1 変数の区分凸最適化問題に帰着できることを示す.

2.1 ユニタリ不変ノルムを用いる場合

問題 (4) における計量にユニタリ不変ノルムを用いる場合, 次の定理で示すように, 変数 X の代わりに X の固有値 λ についての問題に帰着することができる.

定理 1. 問題 (4) における計量をユニタリ不変ノルムとする. 行列 $\widehat{X}$ の固有値分解を $\widehat{X} = P\,\mathrm{Diag}(\widehat{\lambda})P^{\mathrm{T}}$ とし, 次の最適化問題の最適解を λ^* とする:

$$\left|\begin{array}{ll} \text{minimize} & \|\mathrm{Diag}(\lambda - \widehat{\lambda})\| \\ \text{subject to} & \min_i \lambda_i \geq 0, \\ & \max_i \lambda_i \leq \kappa \min_i \lambda_i. \end{array}\right. \tag{5}$$

このとき $X^* = P\,\mathrm{Diag}(\lambda^*)P^{\mathrm{T}}$ は問題 (4) の最適解である.

この定理を示すために 2 つの補題を示す.

補題 2. 任意の $X \in \mathcal{H}^n$ と $\mathcal{H}^n$ 上の任意のユニタリ相似不変ノルム $\|\cdot\|$ に対して次が成り立つ:

$$\|\mathrm{Diag}[\mathrm{diag}(X)]\| \leq \|X\|.$$

証明. X の固有値分解を $X = P\,\mathrm{Diag}(\lambda)P^{\mathrm{H}}$ とすると,

$$\|X\| = \|P\,\mathrm{Diag}(\lambda)P^{\mathrm{H}}\| = \|\mathrm{Diag}(\lambda)\|.$$

ここで $S = P \circ \overline{P}$ とすると, $S \geq O, Se = e, e^{\mathrm{T}}S = e^{\mathrm{T}}$ となるため, S は 2 重確率行列である. したがって, Birkhoff の定理 [5] より, ある置換行列 $\Pi_1, \ldots, \Pi_N$ と $\sum_{i=1}^{n} \alpha_i = 1$ を満たす正の実数 $\alpha_1, \ldots, \alpha_N$

が存在して, $\boldsymbol{S} = \sum_{i=1}^{N} \alpha_i \boldsymbol{\Pi}_i$ とできる. さらに, $\mathrm{diag}(\boldsymbol{X}) = \boldsymbol{S}\boldsymbol{\lambda}$ と 3 角不等式より,

$$\|\mathrm{Diag}[\mathrm{diag}(\boldsymbol{X})]\| = \|\mathrm{Diag}(\boldsymbol{S}\boldsymbol{\lambda})\| = \left\|\mathrm{Diag}\left(\sum_{i=1}^{N} \alpha_i \boldsymbol{\Pi}_i \boldsymbol{\lambda}\right)\right\| = \left\|\sum_{i=1}^{N} \alpha_i \boldsymbol{\Pi}_i \,\mathrm{Diag}(\boldsymbol{\lambda}) \boldsymbol{\Pi}_i^{\mathrm{T}}\right\|$$
$$\leq \sum_{i=1}^{N} \alpha_i \|\boldsymbol{\Pi}_i \,\mathrm{Diag}(\boldsymbol{\lambda}) \boldsymbol{\Pi}_i^{\mathrm{T}}\| = \sum_{i=1}^{N} \alpha_i \|\mathrm{Diag}(\boldsymbol{\lambda})\| = \|\mathrm{Diag}(\boldsymbol{\lambda})\|.$$

したがって,

$$\|\mathrm{Diag}[\mathrm{diag}(\boldsymbol{X})]\| \leq \|\boldsymbol{X}\|.$$

□

補題 3. 任意の $\boldsymbol{X} \in \mathcal{H}^n$ と任意の $i \in \{1, \ldots, n\}$ に対して次が成り立つ:

$$\lambda_{\min}(\boldsymbol{X}) \leq X_{ii} \leq \lambda_{\max}(\boldsymbol{X}).$$

証明. 任意のベクトル $\boldsymbol{v} \in \mathbb{C}^n$ に対して次の不等式が成り立つ:

$$\lambda_{\min}(\boldsymbol{X}) \boldsymbol{v}^{\mathrm{H}} \boldsymbol{v} \leq \boldsymbol{v}^{\mathrm{H}} \boldsymbol{X} \boldsymbol{v} \leq \lambda_{\max}(\boldsymbol{X}) \boldsymbol{v}^{\mathrm{H}} \boldsymbol{v}.$$

ここで $\boldsymbol{v} = \boldsymbol{e}_i$ とすれば示すべき不等式を得る. □

以上の補題を用いて定理 1 を示す.

定理 *1* の証明. 背理法による. 問題 (5) の最適解 $\boldsymbol{\lambda}^*$ に対し, $\boldsymbol{X}^* = \boldsymbol{P}\,\mathrm{Diag}(\boldsymbol{\lambda}^*)\boldsymbol{P}^{\mathrm{T}}$ とする. この $\boldsymbol{X}^*$ が問題 (4) の実行可能解であることは明らか. ここで $\boldsymbol{X}^*$ が問題 (4) の最適解でないとすると,

$$\|\boldsymbol{X}^\dagger - \widehat{\boldsymbol{X}}\| < \|\boldsymbol{X}^* - \widehat{\boldsymbol{X}}\|, \qquad \boldsymbol{X}^\dagger \in \mathcal{S}_+^n, \qquad \lambda_{\max}(\boldsymbol{X}^\dagger) \leq \kappa \lambda_{\min}(\boldsymbol{X}^\dagger)$$

をみたす $\boldsymbol{X}^\dagger$ が存在する. ここで $\boldsymbol{\lambda}^\dagger = \mathrm{diag}(\boldsymbol{P}^{\mathrm{T}} \boldsymbol{X}^\dagger \boldsymbol{P})$ とすると, 補題 3 より,

$$\max_i \lambda_i^\dagger \leq \lambda_{\max}(\boldsymbol{P}^{\mathrm{T}} \boldsymbol{X}^\dagger \boldsymbol{P}) = \lambda_{\max}(\boldsymbol{X}^\dagger) \leq \kappa \lambda_{\min}(\boldsymbol{X}^\dagger) = \kappa \lambda_{\min}(\boldsymbol{P}^{\mathrm{T}} \boldsymbol{X}^\dagger \boldsymbol{P}) \leq \kappa \min_i \lambda_i^\dagger$$

となり, $\boldsymbol{\lambda}^\dagger$ は問題 (5) の実行可能解となる. ところが 補題 2 より,

$$\|\mathrm{Diag}(\boldsymbol{\lambda}^\dagger - \widehat{\boldsymbol{\lambda}})\| = \|\mathrm{Diag}\{\mathrm{diag}[\boldsymbol{P}^{\mathrm{T}}(\boldsymbol{X}^\dagger - \widehat{\boldsymbol{X}})\boldsymbol{P}]\}\| \leq \|\boldsymbol{P}^{\mathrm{T}}(\boldsymbol{X}^\dagger - \widehat{\boldsymbol{X}})\boldsymbol{P}\| = \|\boldsymbol{X}^\dagger - \widehat{\boldsymbol{X}}\|$$
$$< \|\boldsymbol{X}^* - \widehat{\boldsymbol{X}}\| = \|\boldsymbol{P}^{\mathrm{T}}(\boldsymbol{X}^* - \widehat{\boldsymbol{X}})\boldsymbol{P}\| = \|\mathrm{Diag}(\boldsymbol{\lambda}^* - \widehat{\boldsymbol{\lambda}})\|$$

となり, これは $\boldsymbol{X}^*$ が問題 (5) の最適解であることに反する. したがって, $\boldsymbol{X}^*$ は問題 (4) の最適解である. □

2.2 Ky Fan の p-k ノルムを用いる場合

続いて,問題 (4) の計量に Ky Fan の p-k ノルムを用いる場合について考える.このとき,Ky Fan の p-k ノルムの定義より,問題 (5) は次のように変形できる:

$$\left|\begin{array}{ll} \text{minimize} & \sum_{i=1}^{k} |\boldsymbol{\lambda} - \widehat{\boldsymbol{\lambda}}|_{(i)}^{p} \\ \text{subject to} & \min_i \lambda_i \geq 0, \\ & \max_i \lambda_i \leq \kappa \min_i \lambda_i. \end{array}\right. \tag{6}$$

ここで $\boldsymbol{v} \in \mathbb{R}^n$ に対して $|\boldsymbol{v}|$ は $\boldsymbol{v}$ の各成分の絶対値を並べたベクトルを表し,$\boldsymbol{v}_{(i)}$ は $\boldsymbol{v}$ の重複を許して i 番目に大きい成分を表す.この問題は次の定理で示すように 1 変数の凸最適化問題に帰着できる.

定理 4. 問題 (4) において $\boldsymbol{X} = \widehat{\boldsymbol{X}}$ が実行不能であると仮定する.$\widehat{\boldsymbol{X}}$ の固有値分解を $\widehat{\boldsymbol{X}} = \boldsymbol{P}\,\mathrm{Diag}(\widehat{\boldsymbol{\lambda}})\boldsymbol{P}^{\mathrm{T}}$ (ただし $\widehat{\lambda}_1 \leq \cdots \leq \widehat{\lambda}_n$) とする.$f_{p,k,\kappa,\widehat{\lambda}}(\mu)$ を $\mu \geq 0$ に対して $\widehat{\lambda}_l \leq \mu \leq \widehat{\lambda}_{l+1}, \widehat{\lambda}_{u-1} \leq \kappa\mu \leq \widehat{\lambda}_u$ のとき,

$$(\mu - \widehat{\lambda}_1)^p, \ldots, (\mu - \widehat{\lambda}_l)^p, (\widehat{\lambda}_u - \kappa\mu)^p, \ldots, (\widehat{\lambda}_n - \kappa\mu)^p$$

のうち重複を許して大きい方から順に k 個の和を返す関数とし,μ^* を次の最適化問題の最適解とする:

$$\left|\begin{array}{ll} \text{minimize} & f_{p,k,\kappa,\widehat{\lambda}}(\mu) \\ \text{subject to} & \mu \geq 0. \end{array}\right. \tag{7}$$

さらに,次のように $\boldsymbol{\lambda}^*$ を定める:

$$\lambda_1^* = \cdots = \lambda_l^* = \mu^*, \lambda_{l+1}^* = \widehat{\lambda}_{l+1}, \ldots, \lambda_{u-1}^* = \widehat{\lambda}_{u-1}, \lambda_u^* = \cdots = \lambda_n^* = \kappa\mu^*. \tag{8}$$

このとき $\boldsymbol{X}^* = \boldsymbol{P}\,\mathrm{Diag}(\boldsymbol{\lambda}^*)\boldsymbol{P}^{\mathrm{T}}$ は問題 (4) の最適解である.

証明. 定理 1 より,式 (8) で定められる $\boldsymbol{\lambda}^*$ が問題 (6) の最適解であることを示せば十分.問題 (6) に補助変数 μ を導入して次の最適化問題に変形する:

$$\left|\begin{array}{ll} \text{minimize} & \sum_{i=1}^{k} |\boldsymbol{\lambda} - \widehat{\boldsymbol{\lambda}}|_{(i)}^{p} \\ \text{subject to} & \mu \geq 0, \\ & \mu \leq \lambda_i \leq \kappa\mu \quad (i = 1, \ldots, n). \end{array}\right. \tag{9}$$

この問題において μ を固定した次の子問題を考える:

$$\left|\begin{array}{ll} \text{minimize} & \sum_{i=1}^{k} |\boldsymbol{\lambda} - \widehat{\boldsymbol{\lambda}}|_{(i)}^{p} \\ \text{subject to} & \mu \leq \lambda_i \leq \kappa\mu \quad (i = 1, \ldots, n). \end{array}\right. \tag{10}$$

$\widehat{\lambda}_l \leq \mu \leq \widehat{\lambda}_{l+1}, \widehat{\lambda}_{u-1} \leq \kappa\mu \leq \widehat{\lambda}_u$ のときの子問題 (10) を図で表すと図 1 のようになる.図 1 より,次のように定める $\boldsymbol{\lambda}(\mu)$ は $\widehat{\lambda}_l \leq \mu \leq \widehat{\lambda}_{l+1}, \widehat{\lambda}_{u-1} \leq \kappa\mu \leq \widehat{\lambda}_u$ のときの子問題 (10) の最適解であることがわかる:

$$\lambda_1(\mu) = \cdots = \lambda_l(\mu) = \mu, \lambda_{l+1}(\mu) = \widehat{\lambda}_{l+1}, \ldots, \lambda_{u-1}(\mu) = \widehat{\lambda}_{u-1}, \lambda_u(\mu) = \cdots = \lambda_n(\mu) = \kappa\mu.$$

このときの子問題 (10) の最適値 $f_{p,k,\kappa,\widehat{\lambda}}(\mu)$ は,

$$(\mu - \widehat{\lambda}_1)^p, \ldots, (\mu - \widehat{\lambda}_l)^p, (\widehat{\lambda}_u - \kappa\mu)^p, \ldots, (\widehat{\lambda}_n - \kappa\mu)^p$$

のうち重複を許して大きい方から順に k 個の和である. ここで μ の固定を解き, 次の最適化問題を考える:

$$\left| \begin{array}{ll} \text{minimize} & f_{p,k,\kappa,\widehat{\lambda}}(\mu) \\ \text{subject to} & \mu \geq 0. \end{array} \right.$$

この問題の最適解 μ^* とそれに対応する $\lambda(\mu^*) = \boldsymbol{\lambda}^*$ の組は問題 (9) の最適解である. したがって, $\boldsymbol{\lambda}^*$ は問題 (6) の最適解である. □

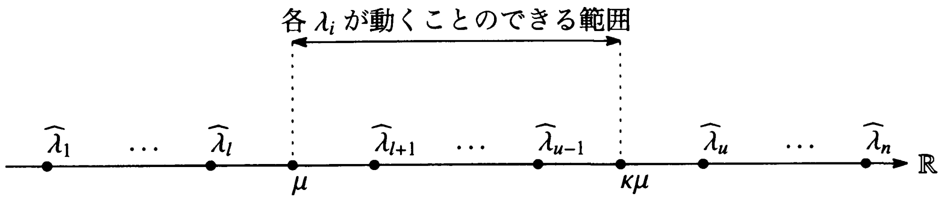

図 1　子問題のイメージ (10)

問題 (7) は 1 変数の区分凸最適化問題である. そのため, 各区分において $f_{p,k,\kappa,\widehat{\lambda}}(\mu)$ の微分係数を計算することで最適解を求めることができる. ここで, 区分の数が高々 $2n+1$ 個であることと, 微分係数の計算が $O(n)$ の計算量で行なえることから, 問題 (7) は $\widehat{\boldsymbol{X}}$ の固有値分解さえ行なっておけば, 2 分探索を用いることで $O(n \log n)$ の計算量で解くことができる.

3　いくつかの場合における解析解

前節では問題 (4) の計量に Ky Fan の p-k ノルムを用いることで 1 変数の区分凸最適化問題 (7) に帰着できることを示した. 本節ではさらに特殊な場合としてスペクトルノルムとトレースノルムを用いる場合の問題 (7) の解析解を示す.

3.1　スペクトルノルムを用いる場合

行列 $\boldsymbol{X} \in \mathbb{C}^{m\times n}$ のスペクトルノルムとは,

$$\|\boldsymbol{X}\|_{p,1} = \sigma_1(\boldsymbol{X})$$

で定義されるノルムで, $k = 1$ とした Ky Fan の p-k ノルムに対応する.

定理 5. 問題 (6) において $k = 1$ とした問題で $\boldsymbol{\lambda} = \widehat{\boldsymbol{\lambda}}$ (ただし $\widehat{\lambda}_1 \leq \cdots \leq \widehat{\lambda}_n$ とする) が実行不能であると仮定する. このとき, 次のように定める μ^* は問題 (7) において $p = 1, k = 1$ とした問題の最適

解である:

$$\mu^* = \frac{1}{\kappa+1}\max\{\widehat{\lambda}_1 + \widehat{\lambda}_n, 0\}.$$

証明. $f_{p,k,\kappa,\widehat{\lambda}}(\mu)$ の定義より,

$$f_{1,1,\kappa,\widehat{\lambda}}(\mu) = \max\{\mu - \widehat{\lambda}_1, \widehat{\lambda}_n - \kappa\mu\}$$

は明らか. $\widehat{\lambda}_1 + \widehat{\lambda}_n < 0$ のとき, $f_{1,1,\kappa,\widehat{\lambda}}(\mu) = \mu - \widehat{\lambda}_1$ より, 問題 (7) の最適解は $\mu^* = 0$. 一方 $\widehat{\lambda}_1 + \widehat{\lambda}_n \geq 0$ のとき,

$$f_{1,1,\kappa,\widehat{\lambda}}(\mu) = \begin{cases} \widehat{\lambda}_n - \kappa\mu & \mu \leq (\widehat{\lambda}_1 + \widehat{\lambda}_n)/(\kappa+1) \text{ のとき}, \\ \mu - \widehat{\lambda}_1 & \text{それ以外} \end{cases}$$

なので, 問題 (7) の最適解は

$$\mu^* = \frac{1}{\kappa+1}(\widehat{\lambda}_1 + \widehat{\lambda}_n).$$

以上より, 問題 (7) の最適解は

$$\mu^* = \frac{1}{\kappa+1}\max\{\widehat{\lambda}_1 + \widehat{\lambda}_n, 0\}.$$

□

3.2 トレースノルムを用いる場合

行列 $\boldsymbol{X} \in \mathbb{C}^{m\times n}$ のトレースノルムとは,

$$\|\boldsymbol{X}\|_{1,n} = \sum_{i=1}^{\max\{m,n\}} \sigma_i(\boldsymbol{X})$$

で定義されるノルムで, $p = 1, k = \max\{m, n\}$ とした Ky Fan の p-k ノルムに対応する.

定理 **6.** 問題 (6) において $p = 1, k = n$ とした問題で $\boldsymbol{\lambda} = \widehat{\boldsymbol{\lambda}}$ (ただし $\widehat{\lambda}_1 \leq \cdots \leq \widehat{\lambda}_n$ とする) が実行不能かつ $n \leq \kappa$ であることを仮定する. このとき, 次のように定める μ^* は問題 (7) において $p = 1$, $k = n$ とした問題の最適解である:

$$\mu^* = \frac{1}{\kappa}\max\{\widehat{\lambda}_n, 0\}.$$

証明. $f_{p,k,\kappa,\widehat{\lambda}}(\mu)$ の定義より, $\widehat{\lambda}_l \leq \mu \leq \widehat{\lambda}_{l+1}, \widehat{\lambda}_{u-1} \leq \kappa\mu \leq \widehat{\lambda}_u$ のとき

$$f_{1,n,\kappa,\widehat{\lambda}}(\mu) = \sum_{i=1}^{l}(\mu - \widehat{\lambda}_i) + \sum_{i=u}^{n}(\widehat{\lambda}_i - \kappa\mu) = [l - \kappa(n-u+1)]\mu + \text{const.}$$

であることは明らか. $\widehat{\lambda}_n < 0$ のとき, $f_{1,n,\kappa,\widehat{\lambda}}(\mu) = n\mu + \text{const.}$ より, 問題 (7) の最適解は $\mu^* = 0$. 一方 $\widehat{\lambda}_n \geq 0$ のとき,

$$f_{1,n,\kappa,\widehat{\lambda}}(\mu) = \begin{cases} (l-\kappa)\mu + \text{const.} & \widehat{\lambda}_l \leq \mu \leq \widehat{\lambda}_{l+1}, \widehat{\lambda}_{n-1} \leq \kappa\mu \leq \widehat{\lambda}_n \text{ のとき}, \\ l\mu + \text{const.} & \text{それ以外} \end{cases}$$

であることと仮定より $l < n \leq \kappa$ であることから, 問題 (7) の最適解は

$$\mu^* = \frac{1}{\kappa}\widehat{\lambda}_n.$$

以上より, 問題 (7) の最適解は

$$\mu^* = \frac{1}{\kappa}\max\{\widehat{\lambda}_n, 0\}.$$

□

トレースノルムを用いる場合の解析ではスペクトルノルムを用いる場合の解析とは異なり $n \leq \kappa$ を仮定する必要がある. しかし, 倍精度浮動小数点演算を行うものとすると, κ は 10^6 程度 (あるいはそれ以上) に設定することが想定されるため, 実用上は大きな問題とならないであろうことを注意しておく.

4 おわりに

本論文で考察した条件数制約つき正定値行列近似問題は対称錐最適化問題として表現することができるが, 本論文ではこの問題を計量に使うノルムを限定することで簡単な構造をもつ非線形最適化問題に帰着できることを示し, いくつかの場合において解析解が得られることを示した.

対称錐最適化問題は確かに多くの問題を表現できる問題のクラスだが, それを解くための計算は必ずしも容易でない. 本論文で示したように, 対称錐最適化問題として定式化された問題は, 問題の構造を利用することで対称錐制約を用いずに表現できることがある. このように, 現実の問題をモデル化する際には解くことができる問題として定式化できた時点で満足するのではなく, より解きやすい問題に変換できないかどうかを考える必要があるだろう.

本論文では最適化問題の解きやすさに着目して議論を行なった. 最適化問題として解きやすいかどうかということと同じかそれ以上に重要な視点は, 最適化モデルとして適切かどうかという観点である. 本論文では, 問題が解きやすくなるという理由でユニタリ相似不変ノルムや Ky Fan の *p-k* ノルムを用いたが, これらを用いることがモデル化として適切かどうか, あるいはどのノルムを用いることが望ましいのかという点について本論文では考察を行なっていない. 今後, これらの観点からの議論が待たれるところである.

参考文献

[1] G. Cornuejols and R. Tütüncü, *Optimization Methods in Finance,* Cambridge University Press, 2007.

[2] M.J. Daniels and R.E. Kass, Shrinkage estimators for covariance matrices, *Biometrics* **57** (2001) 1173–1184.

[3] D.K. Dey and C. Srinivasan, Estimation of a covariance matrix under Stein's loss, *Ann. Statist.* **13** (1985) 1581–1591.

[4] L. Haff, The variational form of certain Bayes estimators, *Ann. Statist.* **19** (1991) 1163–1190.

[5] R.H. Horn and C.R. Johnson, *Matrix Analysis,* Cambridge University Press, 1990.

[6] R.H. Horn and C.R. Johnson, *Topics in Matrix Analysis,* Cambridge University Press, 1994.

[7] W. James and C. Stein, Estimation with quadratic loss, in *Proceedings of Fourth Berkeley Symposium on Mathematical Statistics and Probability,* 1961, pp. 361–379.

[8] O. Leodit and M. Wolf, A well-conditioned estimator for large-dimensional covariance matrices, *J. Multivariate Anal.* **88** (2004) 365–411.

[9] J.H. Won and S.-J. Kim, Maximum likelihood covariance estimation with a condition number constraint, in *Proceedings of Fortieth Asilomar Conference on Signals, Systems, and Computers,* 2006, pp. 1445–1449.